Geospatial Data Spaces

By Martien Vos
27-10-2023
Data spaces require interoperable soft and hard infrastructure

Domains
- Manufacturing
- Green Deal
- Agriculture
- Mobility
- Energy
- Healthcare
- Public Sector
- Finance
- Education

Reference architecture
- Interoperability
  - Data (models & Formats)
  - Data exchange API
  - Provenance & Traceability
  - Data Model (self-explanatory)
- Data Protection
  - Identity & Access Management
  - Usage Control
  - Trusted Exchange (policies/contracts)
- Data value
  - Catalogue (metadata discovery)
  - Clearinghouse (data logging)
  - Vocabularies/Thesaurus/registry
  - Marketplace (app store)

Building blocks
- Governance
  - Use cases
  - Business model/value case
  - Governance model (incl. actor model)
  - Legal Agreement
  - Operational agreement
  - Community Building

Technology Building Blocks
- Technology Infrastructure
- Soft infrastructure
- Hard infrastructure

Data spaces
Traditional department or specific data silos

increasing need to get the **total picture** using multiple data sources
Getting 360° insight out of multiple data sources

Commonly applied method

• Based on data replication
• Needs data ‘refreshment’
• Useful for analytics purposes like data warehouse
• Less useful for federated query due to mismatch and connection issues
• Extremely vulnerable for the possible data model drift on the source side
• ‘Data join’ based on the exact match
Alternative: W3C standard Linked Data (LD)

- No physical data replication
- Real-time data retrieval
- Using the Internet technology
- FAIR data (findability, accessibility, interoperability, and reusability)
- Suitable for federated query
- ‘Data join’ based on meaning (semantic matching)
- Typically using one or more ontologies
  - Set of definitions and rules
- Works also for geographical data through OGC standard of GeoSPARQL
How geospatial data becomes part of Linked Data

I. Define geospatial object in a linked data model like
   • Coordinates
   • term definitions
   • relationships between terms (smart geo objects)

II. Define none spatial objects
   • term definitions
   • relationships between terms

III. Combination of all above

IV. Retrieving geospatial object from linked data by
   • SPARQL
     query terms and relationships
   • GeoSPARQL, special for geospatial objects
     query intersection, overlap, inside etc.
     can also query other none geospatial terms and relationships

Give all objects which are a financial building and have been renovated between date .... (shown on a map)

....and are inside the borders of Madrid (shown on a map)
Example of smart geospatial objects

Via SPARQL query fetching geo-objects and linked objects.

delete power line via SPARQL
Example of smart geo-objects
Geospatial, linked data and domains

- Create or use existing LD models
- Connection across domains
- Created own Geo model (ontology)
  - Annotations at position
  - Smart geospatial-objects
  - Can be linked to other ontologies
- Geospatial applications
- None Geo applications
- Realtime changes on RDF (triples)

RAD applications

Linked Data Models (Ontologies)

Finance domain

Sewage domain

Other domain

Geo domain

Smart Geo-objects

geospatial applications

Delete Line

Sewage pipe material: Concrete

Linked Data Models

PVC

Concrete

€ 5,000
Generate an application based on metadata

Metadata (RDF)

Linked Data

GIS

Linked Data

water

finance

agriculture

GIS Catalogue

Select

Generate

Customer GIS data selection

Metadata for application

Data driven

New mutating GIS application

Linked

Metadata about GIS functionality

9
Technical blocks – data formats – data exchange

REST - GraphQL
GeoSPARQL – SPARQL - SQL
SHACL – JSON schema

Storage layer
- TC
- RDF
- JSON
- Relational
- XML

API layer
- data quality
- Exchange of data (e.g. json-ld)

Access layer

Spatial services
- WMS
- WFS
- Etc
By a Geoserver

3rd party GIS
Developer blocks - data formats - applications

RAD tooling

Developer frameworks

Technical layers

Javascript – Python - Java

REST - GraphQL

SHACL – Json schema

GeoSPARQL – SPARQL - SQL

Storage layer

TC
Long transactions

RDF
Json
Relational
XML

Server site programs
(multi language engine)

API layer and data quality

Access layer

Spatial services
- WMS
- WFS
- Etc
- By a Geoserver

3RD party GIS
Technology infrastructure (cloud example)
Summary

• Linked Data is very suitable for interoperable soft infrastructures
• Linked Data is an alternative method of retrieving relevant data from multiple sources
• It is based on the W3C(www) standard, typically combined with OGC standard of GeoSPARQL for treating geospatial data
• It is a powerful component to make data integration based on the meaning rather than the exact word match.
• Makes data FAIR (findability, accessibility, interoperability, and reusability)
• Unlike traditional data warehouse, it just fetches data relevant to performed query
• A number of implementation platforms for Linked Data combined with GeoSPARQL exist like the one from Oracle
THANK YOU FOR YOUR ATTENTION